GLOBAL DISASTER RISK & RESILIENCE ANALYSIS USING MULTI-SOURCE DATA AND INTERACTIVE VISUAL ANALYTICS

Muhammad Azmat, 23i-2651

Abdullah Ilyas Virk, 23i-2603

FAST-NUCES Islamabad

Abstract-This study performs an integrated analysis of global disaster risk, vulnerability, exposure, and resilience using multi-source datasets including the EM-DAT disaster database, World Bank socioeconomic indicators, and UNDP human development data. After preprocessing, transformation, and normalization, an interactive Tableau dashboard was developed to provide temporal, geographical, socio-economic, and hazard-specific insights. The results reveal contrasting resilience trajectories, strong associations between GDP and resilience indices, and major disparities in disaster outcomes across income groups. The dashboard facilitates exploratory analysis and supports data-informed decision-making.

Index Terms; Disaster Risk, Resilience Analysis, EM-DAT, World Bank Indicators, UNDP HDI, Climate Risk Index (CRI), Tableau Visualization, Socio-Economic Analysis.

INTRODUCTION

Climate-driven disasters continue to intensify, disproportionately impacting low-income and vulnerable regions. Understanding global disaster patterns and resilience requires combining diverse datasets and applying analytical visualization techniques.
This project integrates exposure (E), vulnerability (V), resilience metrics (CRI, DII), and socioeconomic indicators to answer:

1. How has resilience evolved across time and continents?
2. Which countries exhibit high exposure but low vulnerability?
3. Do richer countries recover faster?
4. How does disaster impact vary by disaster type?
5. How do income groups differ in resilience?

An interactive Tableau dashboard was developed to communicate these insights effectively.

OBJECTIVES

The primary objective of this study is to develop an interactive, data-driven analytical system for assessing global disaster risk and resilience across countries. Using multi-source datasets—EM-DAT disaster data, World Bank development indicators, and UNDP human development classifications—the analysis aims to construct standardized resilience metrics and produce exploratory visual analytics.

Specifically, the objectives are:

- To preprocess, integrate, and harmonize heterogeneous datasets related to disaster impacts, socio-economic indicators, and vulnerability/resilience indices.
- To compute derived analytical measures, including Disaster Impact Index (DII), Climate Risk Index (CRI), exposure (E), vulnerability (V), deaths per million, and damage as a percentage of GDP.
- To design and implement interactive visualizations (temporal trends, geographic maps, scatter/bubble plots, and treemaps) enabling exploration across:
    - Time (1950–2022)
    - Geography (countries and regions)
    - Disaster types (earthquake, flood, storm)
    - Socio-economic groups (income and HDI groups)
- To conduct comparative and relational analysis identifying patterns such as:
    - High-exposure but low-vulnerability nations
    - Differences in resilience across income levels
    - Relationship between GDP and resilience metrics
    - Evolution of resilience over time and climate risk trends
- To generate meaningful insights that support data-driven understanding of global disaster resilience and highlight the role of governance, wealth, and development in mitigating disaster impacts.

DATASETS USED

A. EM-DAT: Emergency Events Database (Our World in Data Version)

Source: https://github.com/owid/owid-datasets

Variables used:

- Deaths by hazard type: *earthquake*, *flood*, *storm*
- Total disaster deaths (all disasters)
- Deaths per million
- Affected population (%)
- Total economic damages (USD)
- Damage as % of GDP

Dataset covers 1960–2022 with standardized global reporting.

B. World Bank Core Indicators

Source: https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS

Additional indicators merged through the API:

- Total population
- Urban population (%)

- GDP (constant 2015 USD)
- GDP per capita

Used to compute derived metrics:

- Log GDP
- Exposure-adjusted per capita values
- Normalized population weights for bubble plots

C. UNDP Human Development Dataset (HDR)

Source: https://hdr.undp.org/data-center/documentation-and-downloads

Variables used:

- Human Development Index Group (HDI group)
- Income group classification
- Governance/Institutional capacity classification (when relevant)

Merged using ISO3 country codes.

TOOL & ENVIRONMENT USED

This project was executed using Python for data acquisition, cleaning, preprocessing, and index computation, while Tableau was used for analytical visualization and dashboard development. The following tools, libraries, and computing environment were utilized:

Component	Details
Programming Language	Python 3.12
IDE / Editor Used	VS Code
Data Analysis Libraries	Pandas, NumPy
Data Visualization Libraries (Python)	Matplotlib
Data Format	CSV, Excel (XLSX)
Visualization Platform	Tableau Public

METHODOLOGY

This section outlines the analytical workflow and visualization design implemented to extract resilience insights from the integrated country–year dataset. Each visualization was constructed to support a specific analytical question from R4 and R5.

A. Temporal Analysis (R4.1)

A parameter-driven line chart was developed to analyze the evolution of resilience and disaster-impact metrics from 1960–2022.
Selected Metric (CRI, DII, deaths per million, damage % GDP) was controlled through a Tableau parameter, enabling dynamic metric switching without modifying the worksheet.

Rationale:

- Time series plots are the most effective representation of long-term climatic or socio-economic trends.
- Parameter controls allow exploratory analysis and reduce the need for multiple redundant visualizations.

B. Geographic Comparison (R4.2)

A global choropleth map was constructed to visualize country-level resilience and disaster impacts for any selected year and hazard type. Selected Measure drives the color gradient, and Selected Disaster Type modifies hazard-specific metrics such as deaths.

Rationale:

- Maps offer intuitive interpretation of spatial inequality in resilience.
- Color intensity highlights risk “hotspots,” enabling comparison between regions.
- Detailed tooltips provide context using: CRI, Exposure, Vulnerability, deaths per million, population, GDP, and the current metric.

C. Hazard-Specific Comparison (R4.3)

To compare the severity of earthquakes, floods, storms, and aggregate disaster impacts, the EM-DAT variables were pivoted into a long, tidy format. A bar chart was then constructed to compare total deaths across hazard types for the selected year.

Rationale:

- A pivoted structure allows clean comparison across multiple hazards.
- Bar charts highlight magnitude differences clearly.
- Optional metric switching (damage % GDP, affected population %) enhances interpretability.

D. Socio-Economic Group Comparison (R4.4)

Income groups (High, Upper-Middle, Lower-Middle, Low) were used to compare average CRI, DII, and other selected metrics. A bar chart was used to examine whether higher-income economies demonstrate superior resilience.

Rationale:

- Group-level aggregation exposes structural resilience disparities.
- Socio-economic segmentation aligns with UNDP development categories and supports policy-level insights.

E. Scatter and Bubble Plots for Analytical Storytelling (R5)

Three analytical multi-variable visualizations were developed to address the interpretative questions specified in R5:

1) Exposure vs Vulnerability Scatter Plot

Exposure (E) was plotted against Vulnerability (V), with reference median lines dividing the plane into quadrants (high-high, high-low, low-high, low-low).

Purpose: Identify nations with high exposure but low vulnerability, and highlight those in the high-risk quadrant.

2) GDP vs Resilience Bubble Plot

GDP (log scale) was mapped against CRI, with population as bubble size and income group as color.

Purpose: Evaluate whether national wealth correlates with higher resilience, and observe whether population level moderates the relationship.

3) Treemap of Disaster Impact

A treemap visualizing proportional deaths across hazard types (earthquake, flood, storm, all disasters).

Purpose:

- Show disaster impact distribution at a glance.
- Identify dominant hazard types in total mortality contributions.

F. Dashboard Integration

All analytical views were combined into an interactive Tableau dashboard with global filters (year, disaster type) and parameter controls (metric selection). This enabled exploratory storytelling and met the requirement for multi-dimensional comparative analysis across time, geography, hazard type, and socio-economic segmentation.

VISUALIZATIONS AND ANALYTICAL INSIGHTS

Figure 1. Temporal evolution of CRI and DII across countries (1960–2020).

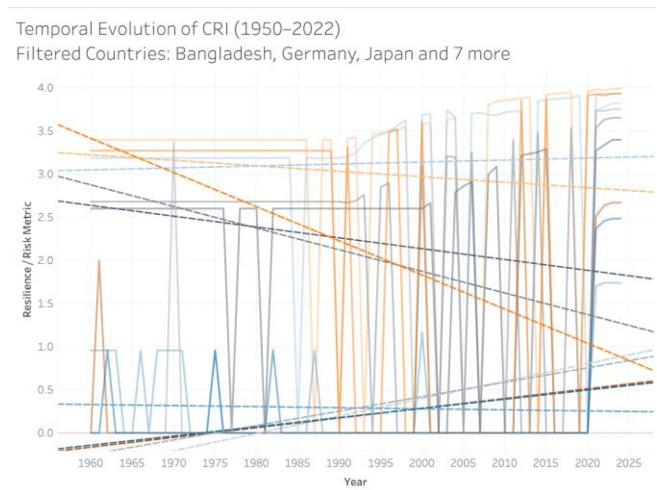


Figure 2. Global choropleth map of disaster impact (damage % GDP).

Figure 3. Disaster-type comparison using pivoted deaths data.

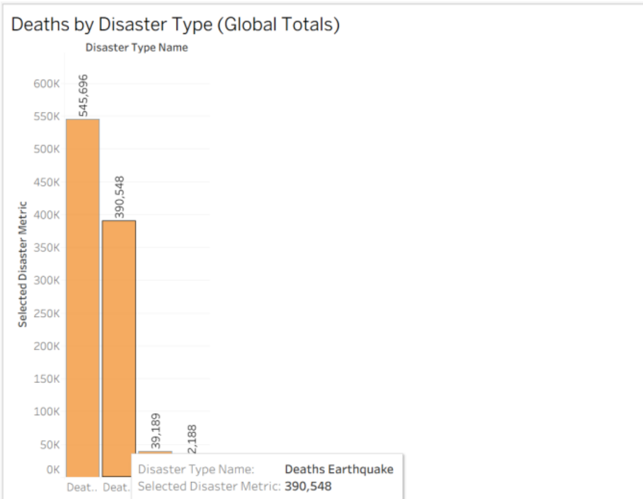


Figure 4. Resilience differences across income groups.

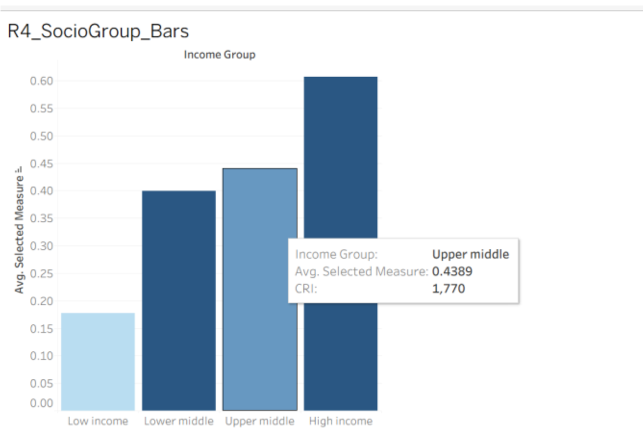


Figure 5. GDP vs. Resilience (Bubble Scatter Plot)

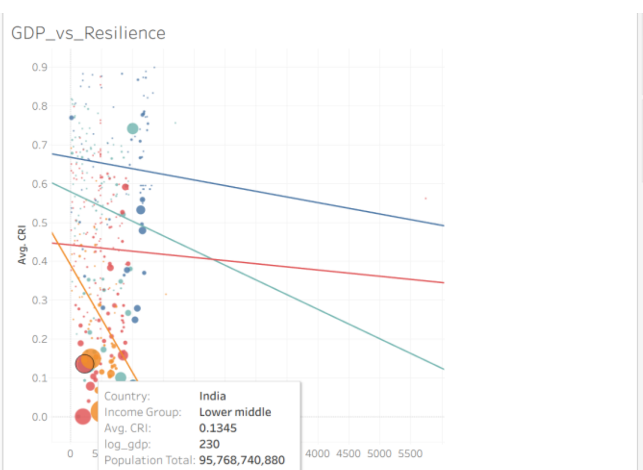
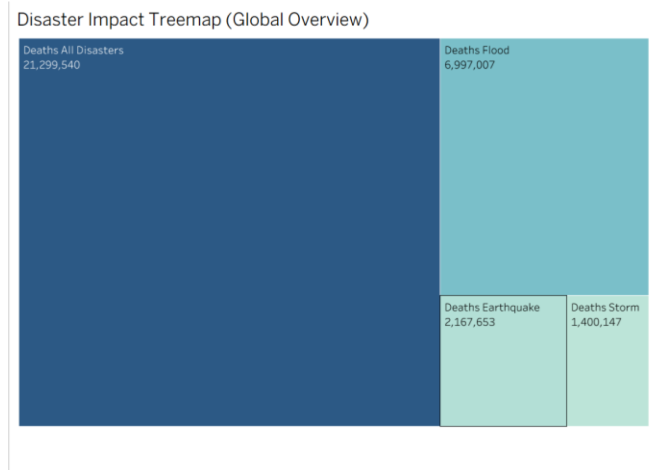


Figure 6. Disaster Impact Treemap (Global Overview)



VISUALIZATION JUSTIFICATION:

Analytical Goal	Visualization Used	Justification
Temporal changes	Line Chart	Best for year-wise evolution
Geographic distribution	Choropleth Map	Shows spatial variation clearly
Compare disaster types	Bar Chart / Treemap	Highlights magnitude differences
Socioeconomic contrast	Grouped Bar Chart	Direct comparison across groups
Risk relationship	Scatter + Reference Lines	Identifies clusters & anomalies
Resilience vs GDP	Bubble Plot	Shows multidimensional correlation

DASHBOARD DESIGN & INTERACTIVITY

The final Tableau dashboard was designed to enable exploratory analysis rather than static reporting. All visual components are interconnected through parameter controls and filters to support multi-dimensional disaster resilience assessment.

A. Parameter Controls

Three primary parameter controls were implemented:

- Choose Measure: dynamically switches between CRI, DII, deaths per million, damage % GDP, affected population %, etc.
- Choose Disaster Type: filters metrics by earthquake, flood, storm, or all disasters.
- Choose Disaster Metric: used in hazard-specific views to toggle between total deaths, damages, and affected population.

These parameters ensure every visualization responds instantly to user-selected analytical focus.

B. Global Filters

- Year Filter (Range slider) applied across all sheets to support temporal comparison.
- Country Filter – enables selecting specific nations for focused analysis when required.

Filters are synchronized across all sheets in the dashboard.

C. Visual Interactivity

Key interactivity features:

- Tooltips enriched with multi-metric information (E, V, CRI, deaths per million, population, GDP).
- Hover-based comparison to instantly view country-level resilience profiles.
- Clickable map regions to update linked visual elements.
- Trendline interaction in scatter plots for quick interpretation of socioeconomic correlations.

D. Dashboard Layout

The dashboard organizes visuals into logically structured analytical components:

1. Temporal View – line graph for resilience evolution over time.
2. Geographical View – choropleth map highlighting global risk disparities.
3. Disaster-Type Comparison – bar charts and treemaps summarizing global impact distribution.
4. Socioeconomic & Analytical Views – scatter and bubble charts exploring relationships (GDP–CRI, Exposure–Vulnerability).

E. User Navigation

Clear titles, legends, and parameter panels guide the user through different analytical questions:

- “How does resilience evolve?”
- “Where are risk hotspots?”
- “Which hazard causes the most impact?”
- “How does wealth relate to resilience?”

This structure ensures simplicity while supporting advanced analytics.

INSIGHTS & FINDINGS

1. Temporal Evolution

- CRI has improved for many high-income nations post-1990.
- Some developing nations show declining resilience due to increased hazard frequency.

2. Geographic Patterns

- Africa and South Asia remain high-risk zones (high vulnerability + low resilience).
- Europe and East Asia show strong resilience scores.

3. Socio-Economic Inequalities

- High-income nations → significantly higher CRI and lower deaths per million.
- Low-income nations → highest exposure and vulnerability.

4. Disaster Type Impact

- Floods account for the largest share of global deaths.
- Storm mortality is high in coastal developing regions.

5. GDP–Resilience Relationship

- Strong positive correlation:
  Higher GDP → higher CRI → better resilience
- Population-heavy nations cluster at lower CRI, even with moderate GDP.

LIMITATIONS

- Missing years in EM-DAT data for several countries.
- HDR income groups change over time but dataset uses static grouping.
- Disaster severity reporting varies by region.

These extensions will strengthen future research, increase model reliability, and support more informed decision-making in clinical practice and public health intervention strategies.

CONCLUSION

This study demonstrates how integrated multi-source datasets can uncover global patterns in disaster impact, resilience, and socioeconomic disparities. The interactive Tableau dashboard allows users to explore country-level risk, compare disaster types, and analyze resilience trends across decades. The results strongly indicate that socioeconomic capacity, especially governance and income, plays a decisive role in shaping resilience.

REFERENCES

EM-DAT Disaster Database (Our World in Data). https://github.com/owid/owid-datasets

World Bank Data Indicators. https://data.worldbank.org

UNDP Human Development Data. https://hdr.undp.org/

ND-GAIN Index Documentation.

Munich-RE Climate Risk Index Methodology.